CAAP Annual Report

Date of Report: *October 7, 2020*

Contract Number: *693JK318500010CAAP*

Prepared for: *U.S. DOT Pipeline and Hazardous Materials Safety Administration*

Project Title: Brain-Inspired Learning Framework to Bridging Information, Uncertainty and Human-Machine Decision-Making for Decoding Variance in Pipeline Computational Models

Prepared by: *North Dakota State University*

Contact Information: *Ms. Zi Zhang, PhD student, Email: zi.zhang@ndsu.edu@ndsu.edu, Phone: 701-231-7204; Mr. Xingyu Wang, PhD student, Email: xingyu.wang@ndsu.edu, Phone: 701-231-7204; Mr. Matthew Pearson, M.S. student, Email: matthew.pearson@ndsu.edu, Phone: 701-231-7204; Ms. Li Shang, PhD student, Email: li.shang@ndsu.edu; Dr. Hong Pan, Postdoc, Email: hong.pan@ndsu.edu; Dr. Zhibin Lin, Email:zhibin.lin@ndsu.edu, Phone: 717-231-7204*

For quarterly period ending: *October 7, 2020*

Business and Activity Section

(a) Generated Commitments

No-cost extension due to COVID-19 pandemic

Some purchase of steel plates and piezoelectric sensors

(b) Status Update of Past Quarter Activities

The research activities in the annual report 2 included: (i) Task 4: Completed efforts on decoding uncertainty due to measurement noise, operational condition (temperature variances), and structural initial nonlinearity; aging effects toward data variances are in progress. (ii) Task 5: Completed efforts on decoding heterogeneous data sets; (iii) Task 6: Development of learning for human-in-the-loop decision making, semi-automated human-machine decision making, and decision-tree based human-machine decision making. Survey and method calibration are in progress.

(c) Cost share activity

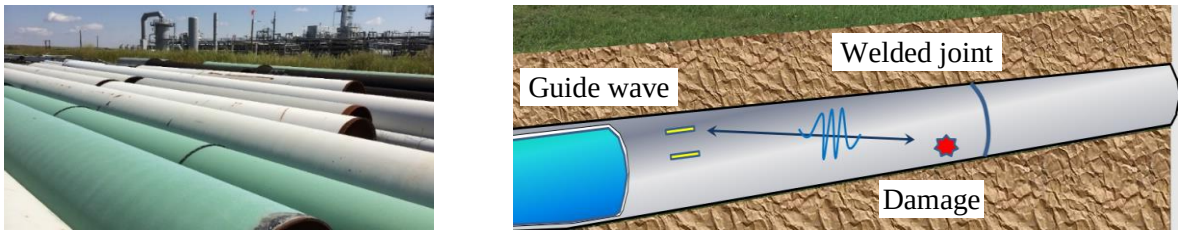
Cost share was from the graduate students' tuition waiver.

(d) Summary of detailed work for Tasks 4-6

Task 4: Decode Variance from Uncertainties

This task aimed to address the challenge to decode variance due to various uncertainties associated to measurement noise, operational, and nonlinearity due to material discontinuity.

As for large volumes of ultrasonic data, numerous uncertainties and complex guided wave propagation in oil/gas pipeline, as schematically in **Fig. 1**, feature extraction may cause the limitation. However, deep learning methods, such as convolutional neural networks (CNN), was used herein to enhance information extraction and better classify structural uncertainty of from data in pipeline associated with a high levels of variances, including measurement noise, nonlinearity and other uncertainty. Clearly, oil/gas pipeline structures are often exposed to complex environment with high levels of uncertainty, as schematically shown in **Fig. 1b**. As a result, Lamb wave signals collected from complex structural systems in fields could be highly affected by structural uncertainty, which in turn affects the effectiveness of the methods for engineering applications. We discussed the simulation model of the pipeline and several states were designed to detect the damage using CNN algorithm. As stated in last report, the framework of this study was shown in **Fig. 2**. The study was presented herein to address the effectiveness of the proposed deep learning methods when handling structural uncertainty due to noise level and material discontinuity from weldment that pipe engineers often face with in field.



(a) Pipes used for oil/gas transmission line (b) Schematics of guided wave along a pipeline

Fig. 1 Data collected from guided wave along a pipeline

The merits of using the proposed learning framework over conventional physics-based signal process were mainly on:

(a) Handling nonlinear and high-dimensional features; physics-based features such as amplitude, phase change, and correlation coefficient, which are often used explicitly for determining damage level and size, could be insensitive to defects in some cases when facing with complexity of guided wave multimodal interaction, noise or other interference. Differently, deep learning could automatically extract sensitive features related to structural and material discontinuity, with less physical representation.

(b) Tackling more structural complexity with less physical restraints; Guided wave exhibits non-stationary and nonlinear behavior, experiencing complex dispersion and coherent multi-mode interaction. Different to physics-based methods that attempt decomposition of mixed modes for signal process, the machine learning could extract sensitive damage features, with less or without such physical restraints. As a result, with representative data, the machine learning could provide better damage detection with minimized explicit formation that physics-based methods highly rely on.

(c) Uncovering structural uncertainty; Consider that oil/gas pipelines are often exposed to high levels of uncertainty, structural uncertainty is one of challenges for physics-based methods. The cases were designed in this study to address this challenge and demonstrate the effectiveness of the proposed learning framework under structural uncertainty due to noise level and material discontinuity from weldment. The findings were expected to provide new vision using machine learning methods for pipeline engineering applications.

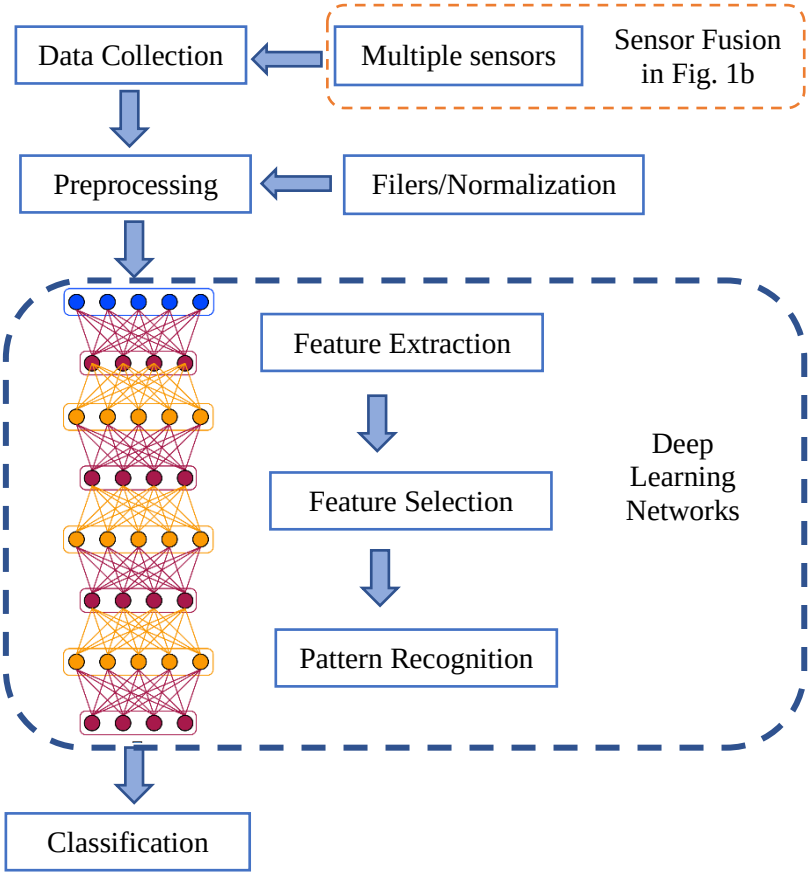
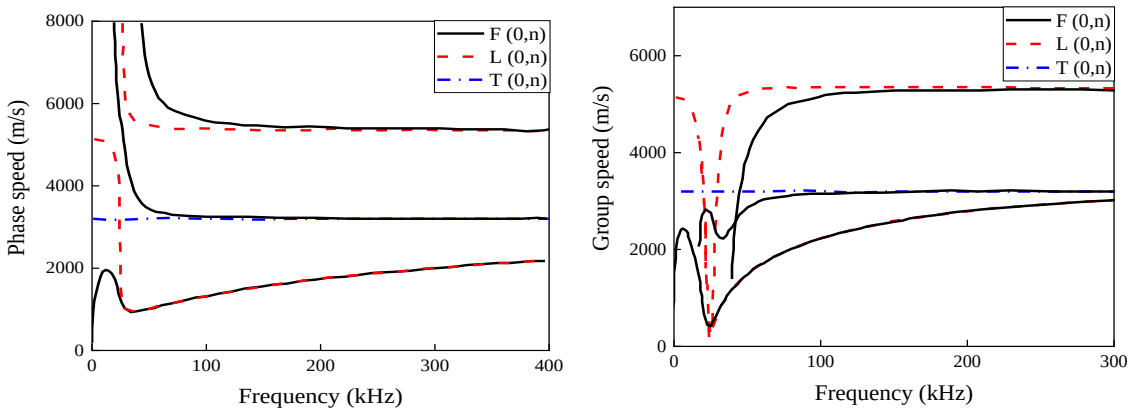


Fig. 2 Framework in this study

Oil/gas pipeline was simulated using 3D FE modeling through COMSOL. The prototype of a steel pipeline was selected from the literature, where its dimension is 76-mm in outside diameter and 4 mm in wall thickness, and with a length of 2000 mm.

(a) Ultrasonic guided waves

Ultrasonic guided waves generated by piezo actuators can propagate along the pipe and change when excited guided wave signals encounter defects or other material discontinuity. Different damage location and severity can cause the wave scattering in the form of mode conversion, reflection and transmission which makes the ultrasonic guided wave-based damage detection complicated. As a result, acquiring the knowledge of guided wave dispersive behavior, mode shapes, and suitable frequency range will make the damage detection susceptible to control. According to the **Fig. 3**, L (0, 2) mode in range 50 to 150kHz has lower dispersion, higher speed and lower distorted which is commonly used in testing. As such, L (0,2) mode was select to detect damage in this research.



(a) Phase speed (b) Group speed

Fig. 3 Multiple mode of guided wave

(b) Impacts of diameter of pipe on signal characteristics

Diameter of pipeline varies from less than one inches to relatively an important factor for guided wave testing. In experiment, when same actuators glued around the pipeline, the big diameter need more energy to excite. In addition, the propagation paths of guided waves are changed in big diameter pipes. In this section, 4 different diameter pipelines were designed to investigate the guided wave propagated in different pipes, including 3 inches, 6 incher, 12 inches and 18 inches, shown in **Table 1**. The length of the pipes was 20 times of the 3 inches (1520 mm). The thickness was 4 mm. A though-thickness notch shaped damage where located at the middle of the pipe. The width of the notch was 1 mm, and the length of the notch was depended on the diameter of the pipe, which was 0.01 time of the diameter. The model was shown in **Fig. 4**. Totally, 16 excitation nodes were located at the center of the left boundary. The excitation wave was 100 kHz 5-cycle sine function operated with a Hanning window. The received signals were normalized and illustrated in **Fig. 5**. Clearly, with the pipe diameter increase, the positions of the reflections were same. However, the amplitude of the boundary reflection was reduced, and the damage reflection was different.

Table 1 Test matrix for computation modeling

Case	Label	Diameter of pipeline	Damage type	location	Noise Interference
Variance due to diameter	State #1	3 inches	Notch	middle	Noise levels of from 60 dB to 100 dB
	State #2	6 inches	Notch	middle	
	State #3	12 inches	Notch	middle	
	State #4	18 inches	Notch	middle	

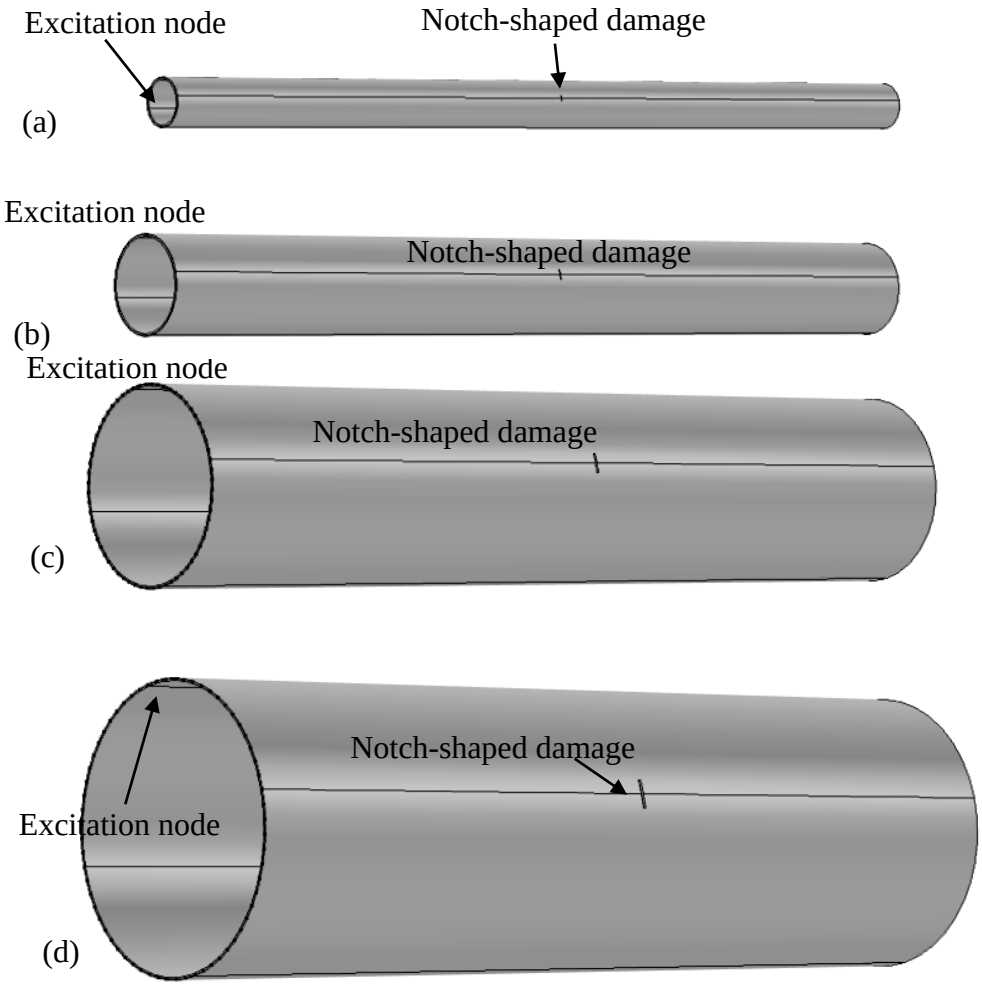


Fig. 4 Pipe model with different diameters, ranging from 3 in. to 18 in.

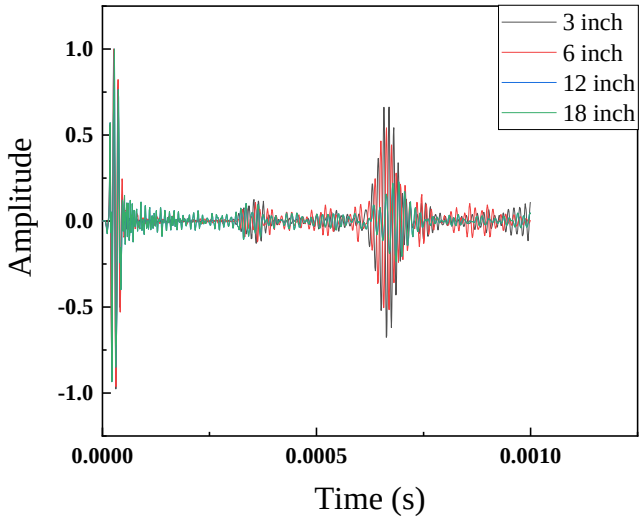


Fig. 5 Received signals

As such, the results herein confirmed the signal characteristics of pipes under varying diameters exhibited identical trend. The proposed learning framework could effectively capture the differences due to changed diameters. Therefore, for simplicity, we used 3-in. diameter pipe as our representatives for data fusion (in both numerical and experimental studies). Note that we will conduct 6-in. diameter pipes as well in the coming periods to further confirm this statement to ensure the effectiveness of the proposed learning for different scenarios.

Sub-Task 4.1. Reducing variance due to measurement noise

To address the impacts of noise levels to signals, the overall of 15 different states were designed and listed in **Table 2**, including three different scenarios designed with notch-shaped damage (i.e., different damage location, size and depth).

To consider the uncertainty happened in actual situation, noise was added to the collected signals based on the signal to noise ratio (SNR) that represents the ratio of the signal strength to the background noise strength as:

$$SNR_{dB} = 10\log_{10}\left(\frac{P_{\text{signal}}}{P_{\text{noise}}}\right) \quad (1)$$

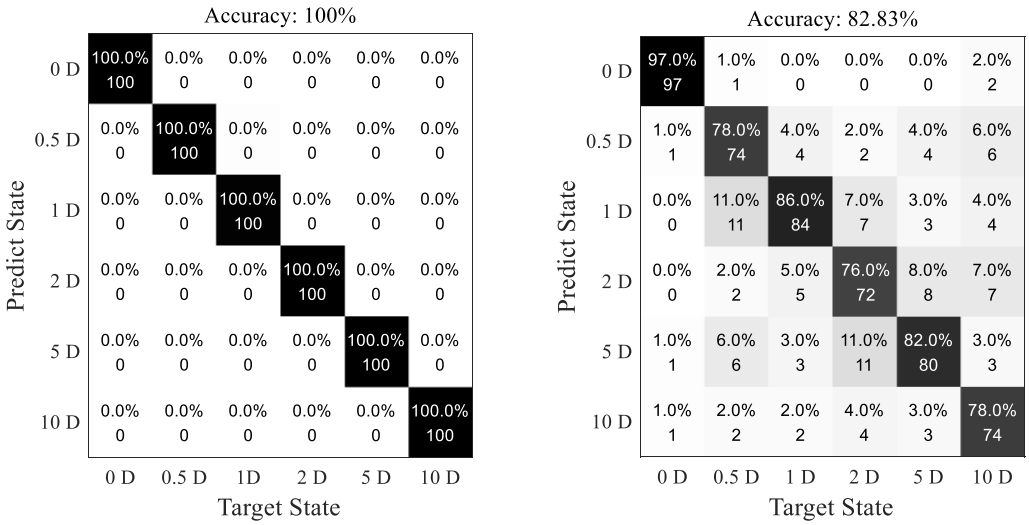
where P_{signal} and P_{noise} are the average power of signal and noise by the dB scale, respectively. Five different noise levels, ranging from 60 dB to 100 dB, were selected for machine learning to check the sensitivity of the uncertainty due to noise.

Table 2 Test matrix for computation modeling

Case	Label	Damage location	Damage size	Damage depth	Noise Interference
Reference	State #1	/	/	/	Noise levels of from 60dB to 100dB
Variance due to damage location	State #2	0.5* D_{out}	0.1* D_{out}	4 mm	
	State #3	1* D_{out}	0.1* D_{out}	4 mm	
	State #4	2* D_{out}	0.1* D_{out}	4 mm	
	State #5	5* D_{out}	0.1* D_{out}	4 mm	
	State #6	10* D_{out}	0.1* D_{out}	4 mm	
Variance due to damage size	State #5	5* D_{out}	0.1* D_{out}	4 mm	
	State #7	5* D_{out}	0.2* D_{out}	4 mm	
	State #8	5* D_{out}	0.3* D_{out}	4 mm	

Variance due to damage depth	State #9	5* D_{out}	0.4* D_{out}	4 mm
	State #10	5* D_{out}	0.5* D_{out}	4 mm
	State #11	5* D_{out}	0.6* D_{out}	4 mm
	State #12	5* D_{out}	0.7* D_{out}	4 mm
	State #13	5* D_{out}	0.1* D_{out}	1 mm
	State #14	5* D_{out}	0.1* D_{out}	2 mm
	State #15	5* D_{out}	0.1* D_{out}	3 mm
	State #5	5* D_{out}	0.1* D_{out}	4 mm

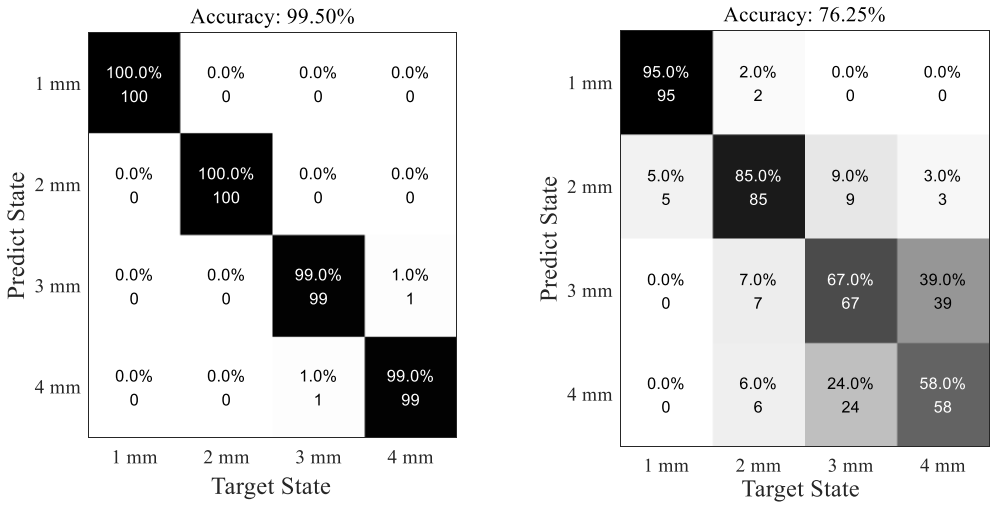
Validating the classification model by testing data, the results were illustrated in **Fig. 6**. 600 signals were input into the pretrained model, and the prediction result of each data was directly obtained. Obviously, the CNN method trained by a part of the signals could accurately identify the damage location in most situations, which acquired 100% classification under noise level of 100 dB to 70dB. **Fig. 6(a)** showed confusion matrix of the testing result in 70 dB. However, the special case was happened at 60 dB, which the accuracy of the test was decreased dramatically to 82.83%. The data in each category was misled into other categories to some extent. Specifically, 11 of the signals whose damage was located at 0.5 D_{out} from the weldment were misclassified as group 1 D, and 6 of them was misled into 5 D_{out}. The accuracy of the fourth class was only 76%, which means the prediction of 24 test signals in this class were failed. The reason of the low accuracy was that the high level of noise added to the signals replaced the most of original features with the noise features. Therefore, the classification method was ineffective when the training data was Indistinguishable.



(a) SNR = 70 dB (b) SNR = 60 dB

Fig. 6 Testing results

The testing results of the damage depth were shown in **Fig. 7**, when the SNRs were 70 dB and 60 dB. As noise level was 70 dB, 2 out of the 400 testing points were incorrect which made the accuracy rate equal to 99.5%. Similar with the situation in damage location and damage size, when the SNR reached to 60 dB the testing accuracy was dropped down dramatically, only 76.25%. The errors were mainly occurred at the damages belonging to 3-mm depth and 4-mm depth state whose accuracy were 67% and 58% respectively.



(a) SNR = 70 dB (b) SNR = 60 dB

Fig.7 Testing results of damage depth

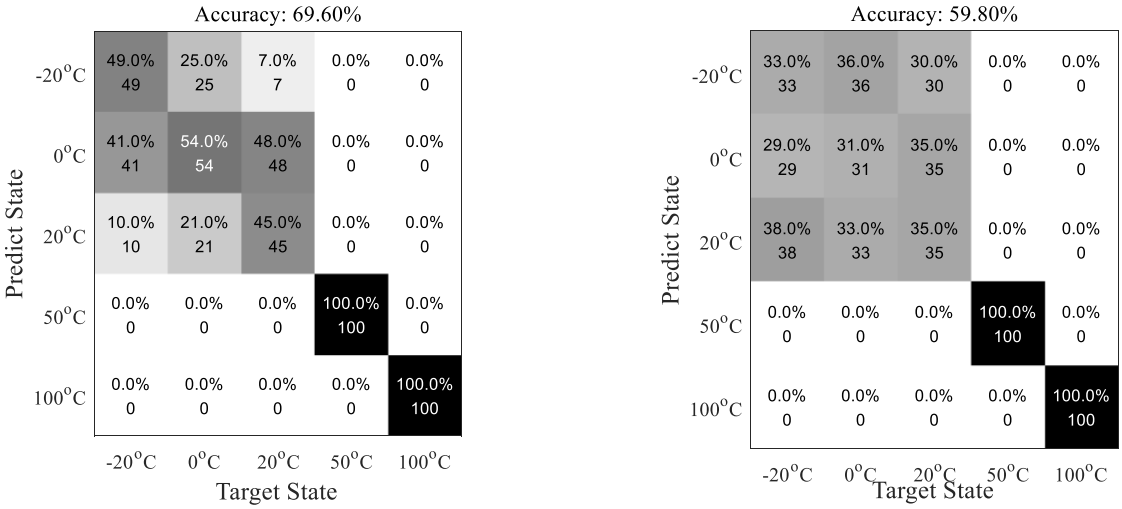
Sub-Task 4.2. Reducing variance due to data under changed operational environments

In this study, the temperature change was treated as one of operational variance and elucidate its impacts on the data classification. As listed in **Table 3**, scenarios were designed to account for variance in terms of temperature change, ranging from -20 ^{o}C to 100 ^{o}C, which pipeline measurement could experience under operational conditions. The weld with and without defect was considered as the case for determining of data classification, where Defect 1 denotes lack of fusion. Note that uniformly distributed temperature was ideally assumed through the entire pipe in this study. However, an entire pipe often experiences a high level of temperature deviation, particularly at the locations of weldment and damage areas, which will be investigated in the future study.

Table 3 Test matrix for computation modeling

Case	Label	Welding defect type	Severity of Welding defect	Temperature	Noise Interference
Reference	State #1	/	/	20^{o}C	Noise levels of from 60dB to 100dB
Variance due to temperature	State #2	Defect 1	1%	-20^{o}C	
	State #3	Defect 1	1%	0^{o}C	
	State #4	Defect 1	1%	20^{o}C	
	State #5	Defect 1	1%	50^{o}C	
	State #6	Defect 1	1%	100^{o}C	

Figs. 8(a)-8(b) were plotted for the confusion matrix of classification with temperature changes under different noise levels. Under high temperature, the prediction was 100%. The accuracy reduced to 69.6% when SNR was 80 dB, as shown in **Fig. 8(a)**. The prediction among -20^{o}C, 0^{o}C and 20^{o}C was terrible, as 49%, 54% and 45%, respectively. On the contrary, the result in high temperature was 100%. When the noise level was about 70 dB, the accuracy of prediction dropped to 60%, and similar to the case under SNR=80 dB, large portions among -20^{o}C, 0^{o}C and 20^{o}C had a high misleading, as shown in **Fig. 8(b)**.



(a) SNR = 80 dB (b) SNR = 70 dB

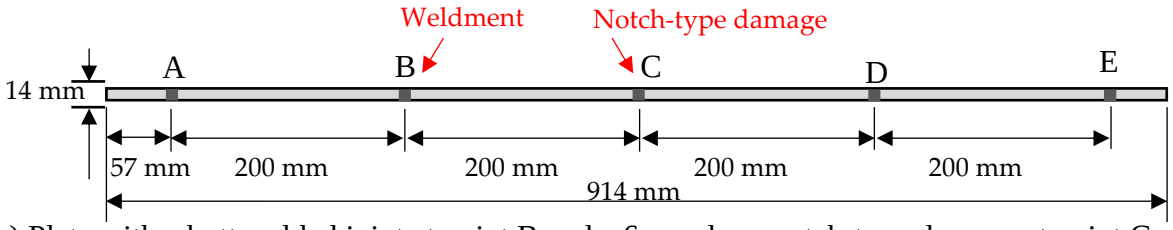
Fig. 8 Confusion matrix under different noise levels

Sub-Task 4.3. Reducing variance due to structural initial nonlinearity

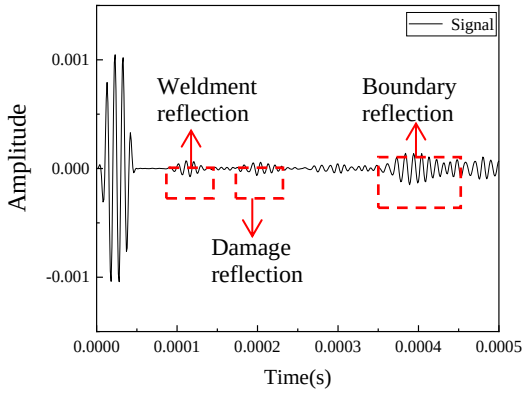
Pipeline systems exhibit highly initial nonlinearity due to aging and material degradation. Initial nonlinearity has to be used in feature pools as baseline. In this study, we focused on investigation of variances due to weldment that is one of structural initial nonlinearity and addressed their impacts on classification.

Material discontinuity due to weldment creates more complexity for lamb wave signal process. This section was to discuss the effectiveness of the proposed method for classifying such structural uncertainty. To test the accuracy of this model, a new dataset was built by numerical simulation method. The identical plate, illustrated in Fig. 4, was modified by adding a butt weldment at location of point B and a 6mm-long notch at location of point C, shown in **Fig. 9(a).** The width of the weldment was 5 mm and the welding filler was Ti-6Al-4V. To enlarge the data, 175 signals were augmented using white Gaussian noise with different levels.

With the interaction of the weldment, the received signal had more reflected packages than that of the previous one, shown in **Fig. 9(b)**. From the signal, the reflections came from the weldment, damage and the boundary. The label of each data was predicted by SVM which was trained in section 5. **Table 4** showed the comparation of the predicted result in 6 mm-long damage and the one added the weldment. Clearly, the prediction of the damage with the weldment post challenge in classification as compared to cases without weldment. Specifically, in most of the case (shown in **Fig. 10**), 6-mm long damage was classified accurately which was 100% expect the SNR equal to 80 dB. With the weldment appeared, it was interfered the signal and reduced the accuracy of the prediction. However, most of the damage could be test by this model. A total of 80% of damaged cases were classified into 6 mm-long damage group, and 20% of the damage was predicted as 4 mm-long at 100 dB. The misleading was increased to 26.9% when noise level approached 90 dB.



(a) Plate with a butt welded joint at point B and a 6-mm long notch-type damage at point C.

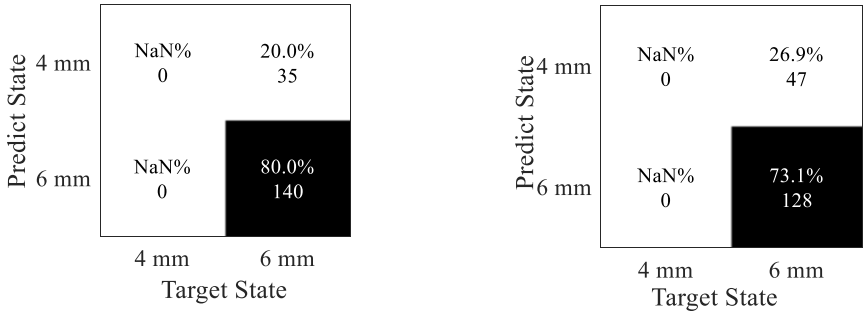


(b) Signal collected from point A

Fig. 9 Plate with weldment and 6 mm-long notch-type damage

Table 4 Prediction

Noise level	120 *dB*	110 *dB*	100 *dB*	90 *dB*	80 *dB*
Without weldment	100.00%	100.00%	100.00%	100.00%	56.4%
With weldment	100.00%	100.00%	80.00%	73.10%	65.71%



(a) SNR = 100 *dB*, accuracy = 80.0% (b) SNR = 90 *dB*, accuracy = 73.1%

Fig. 10 Classification of damage sizes under two noise levels

Task 5: Decode Variance from Heterogeneous Data Sets

This task aimed to address the challenge due to various heterogeneous data sets and unveil the fundamental nature of data types.

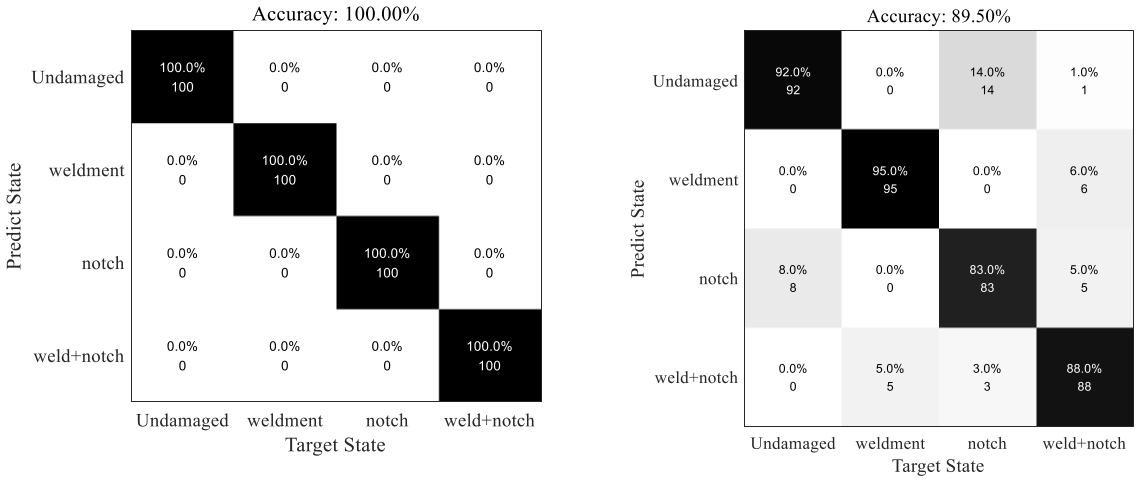
Sub-Task 5.1. Mixed fusion method for heterogeneous data sets

First attempt was to category the multiple type and understand how to mitigate the influence of different contamination. To accomplish this task, we used a mixed data fusion strategy to clear and assemble data. This mixed data fusion process including sensor level fusion and central-level fusion. The final result of this data fusion process is united to three different formats based on the original signal's format and the application purpose.

Sub-Task 5.2. Data classification between image and sensory data

The second attempt was made using common images and sensory data. The results of the damage detection were 4 different signals from receivers located circumferentially around the pipe. Considering the interrelationship between signals from different locations, four signals were combined as a matrix input into the deep learning.

Fig. 7 represented the confusion matrix of the testing data in SNR equal to 70 dB and 60 dB respectively. When the noise level is lower than 70 dB, the accuracy of the prediction for the testing data was 100%. However, in 60 dB condition, the accuracy of the identification was 89.5%.



(a) SNR = 70 dB (b) SNR = 60 dB

Fig. 7 Confusion matrix for time-series data classifcation

Task 6: Explore Variance from Human-Machine Interfaces

This task aimed to explore variance due to human-machine interfaces and mixed human-machine data/information

Sub-Task 6.1. Physics based human interpretation of data vs. Semi-automated vs. Automated machine interpretation of data to eliminate dependency to operators

6.1.1 Physics based human interpretation of data

Take mechanical damage description if measurement is based on guided wave for an example, the physics-based methods are based on physical characteristics of guided waves along pipeline systems. The captured sensor data are usually integrated in the analytical models or simulations for calibration or interpretation based on expert experience to identify physical characteristics and assess the conditions of a pipe. Analytical models and simulation techniques are now well-established, which are still the major base for experts in the pipeline community.

Lamb wave exhibit apparently non-stationary and nonlinear behavior. Mechanical damage description could be interpreted under frequency-, time-, or time-frequency-domains. In time domain, physics-based features play an important role in Lamb wave feature extraction. Amplitude, energy, and correlation coefficient are three features which can represent the wave characteristic. The amplitude was obtained by the peak value of the damage wave packet. The energy calculated by the root mean square of wave (RMS) in the damage part were defined as

$$\mathrm{RMS} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} e_i^2} \tag{2}$$

where n is the number of data point and e_i is the signal. The correlation coefficient under the damage state was used to compare with that of the health state. In frequency domain, the amplitude was extracted as the features.

6.1.2 Physics-guided semi-automated interpretation of data

The physics-guided semi-automated interpretation of data are often based on critical physical characteristics of guided waves under different domains. As shown in **Fig. 8(a)**, the shallow learning is often employed for such a purpose. From the point of view of learning architecture, the shallow learning, such as the SVM or conventional artificial neural network, uses one or zero hidden layer for a shallow

linear pattern separation. The shallow learning highly relies on the quality of the hand-crafted features (physics-guided features), suitable for well-constrained cases (such as physical features, such as amplitude, frequency or RMS)

Take SVM as example for demonstration of the semi-automated interpretation. Consider the set of training vectors$(\mathbf{x}_1, y_1)$,… $(\mathbf{x}_k, y_k), \in \mathbf{R}^N$ belonging to two classes ($y_i = \{-1, 1\}$). The aim is to look for the hyperplane to separate the data:

$$(\mathbf{w} \cdot \mathbf{x}) + b = 0, \mathbf{w} \in \mathbf{R}^N, b \in \mathbf{R} \tag{3}$$

where $\mathbf{w}$ is the weight parameter controlling the orientation of the hyperplane; b is a scalar threshold adjusting the bias of margins between the optimal hyperplane and the support vectors. Then the feature space for the linear classifier is shown

$$f(\mathbf{X}) = \text{sgn}((\mathbf{w} \cdot \mathbf{x}) + b) \tag{4}$$

For the simplest case of a two-dimensional space, several linear classifiers could separate the data. The goal is to look for the hyperplane with largest margin, which is called the optimal hyperplane. Thus, all the training data are satisfying the constraints as follows

$$\mathbf{x}_i \cdot \mathbf{w} + b \geq +1 \text{ for } y_i = +1 \tag{5}$$

$$\mathbf{x}_i \cdot \mathbf{w} + b \leq -1 \text{ for } y_i = -1 \tag{6}$$

The geometric distance from data point to hyperplane ($\mathbf{w}$, b) is shown

$$d\big((\mathbf{w}, b), \mathbf{x}_i\big) = \frac{y_i(\mathbf{x}_i \cdot \mathbf{w} + b)}{\|\mathbf{w}\|} \geq \frac{1}{\|\mathbf{w}\|} \tag{7}$$

To obtain the optimal hyperplane, the maximum distance to the closest data points should be find. From Eqn. (13), acquiring the maximum distance is same as finding the minimum value of $\|\mathbf{w}\|$. Therefore, the optimization could also change into a convex quadratic programming problem [41]

$$\text{Minimize } \boldsymbol{\Phi}(\boldsymbol{w}) = \frac{1}{2}\|\boldsymbol{w}\|^2 \tag{8}$$

Lagrange multiplier is the main method to finding the local maxima and minima of a function subject to equality constraints. The problem is transformed into [41]

$$L(\mathbf{w}, b, \Lambda) = \frac{1}{2}\|\boldsymbol{w}\|^2 - \sum_{i=1}^{k} \lambda_i [y_i(\mathbf{w}^{\mathrm{T}}\mathbf{x}_i + b) - 1] \tag{9}$$

where $\Lambda = (\lambda_1 \cdots \lambda_k)^{\mathrm{T}}$ are the Lagrange multiplier. The $L(\mathbf{w}, b, \Lambda)$ has to be minimized with respect to $\mathbf{w}$ and b, and maximized with respect to $\Lambda \geq 0$.

The decision function is given by

$$f(\mathbf{X}) = \text{sgn}(\textstyle\sum_{i=1}^{k} \lambda_i^* y_i K(x, x_i) + b) \tag{10}$$

where the $K(x, x_{\mathrm{i}})$ is the kernel function, and three commonly used types are gaussian radial basis function (RBF), polynomial function and sigmoid function. In this paper, RBF was selected as the kernel function.

$$K(x, x_{\mathrm{i}}) = \exp(-\gamma \|x_i - x_i\|^2), \quad \gamma > 0 \tag{11}$$

In general, the kernel function, shown in Eqns. (11), tend to construct a higher dimensional feature space and allows a projectile of data to this hyperplane to achieve being linearly separable. The kernel function helps SVM much more suitable for different dataset which can be used in non-linear classification. The different kernel functions have their applicability, including computation cost and parameter tuning. To enhance the accuracy of the damage prediction, it is important to select suitable penalty coefficient and kernel function parameter for the SVMs.

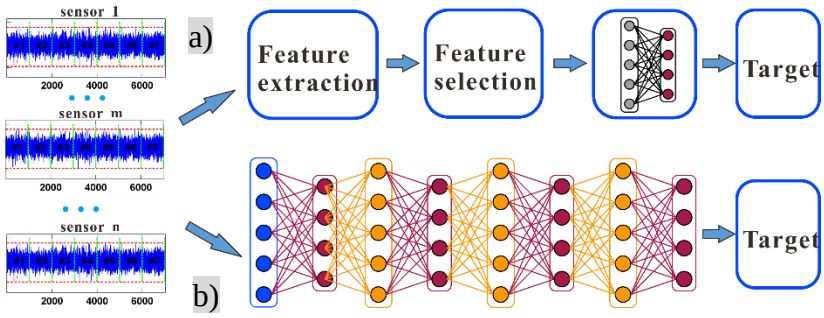


Fig. 8 Schematics of (a) Semi-automated interpration of data using shallow learning (i.e., SVM); and (b) Full automated machine interpration of data using deep learning

6.1.3 Automated machine interpretation of data to eliminate dependency to operators

As shown in **Fig. 8(b)**, automated machine interpretation of data is often achieved using deep learning that stems originally from artificial neural networks with one shallow hidden layer architecture. The deep learnings are constructed by stacking multiple layers in hierarchical architectures, as illustrated in **Fig. 8(b)**. The deep architectures exhibit their merits over their counterparts (e.g., shallow learning), as they attempt to find statistical representation based on end-to-end decoding without applying any preprocessing (predetermined features) as well as no applying any post-processing (feature extraction and/or feature selection).

6.1.4 Comparative study of physics-based, semi-automated, and full automated interpretation of data

A prototype of a thin narrow-strip plate was selected from the literature work. The plate had a dimension of 914 mm by 14 mm with a thickness of 1.6 mm. The piezo actuator was installed in front of the beam, as shown **Fig. 9**. Five different spots, *A* to *E*, were selected to receive the signal, as shown in **Fig. 9**. The damage was located at the point C (457mm away from left side) of the plate. An 8-mm through-the-thickness notch was defined in the COMSOL to simulate the shape of the damage.

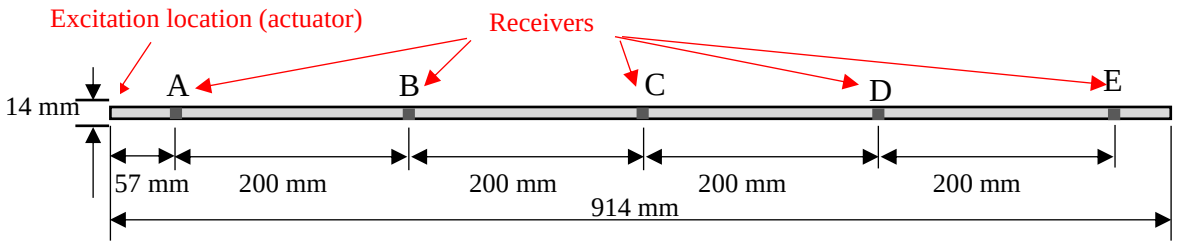


Fig. 9 The thin plate with mechanical damage

Although the physics-based features can classify the damage into different state at a low noisy environment (SNR=100 dB), it is hard to get a high accuracy result under higher noise level. **Table 5** shows the result of the classification with different method and features. Three physics-based features were used for classification by the traditional way respectively. Then, SVM was involved by three feature groups, including physics-based features (Amp, Frq, Cor and RMS), all feature pools (4 of the physics-based features and 12 of the wavelet coefficients) and the selected features by feature selection methods.

By traditional method, Amp, Frq and RMS presented the good result at 120 dB and 110 dB. However, with the level of SNR increased, the accuracy of the separation was dropped down sharply. Specifically, only 39.43% of the data can be classified correctly through RMS which was the highest one comparing with the Amp (19.43%) and Frq (34.86%). On the other hand, SVM method showed the superiority by the high dimensional features which was much more accurate especially at high noise level states. In this method, the results were distinct by different feature groups. Using physics-based features to train the data, although the accuracy was reached to 100% (SNR = 120 dB), the ratio began to reduce into 98.86% at 110 dB and then decreased to 53.17% at 80 dB. To increase the dimension of the features, all the features were used for training data. The accuracy of each state was not increased dramatically, which was lower than the model trained by selected features. Clearly, using selected features, 95.43% of the data was identified at 100 dB comparing with the 84% by all the features. In the case of SNR equal to 80 dB, nearly 17% of the accuracy was increased by feature selection. Therefore, physics-guided semi-automated interpretation of data using SVM combining with the feature selection method can increase the accuracy of the classification, as compared to traditional physics-based models. In addition, with the increase of uncertainty, such as measure noise reaching up to 90 or 80 dB, the semi-automated methods could suffer from significant reduction in accuracy. The relatively low accuracy (such as 56%) could provide high risk in false alarm and lead to wrong decision making.

Table 5 Accuracy of different features

Method	**Classification by physics-based**			**Physics-guided semi-automated using SVM**	
Features	*Amp*	*Frq*	*RMS*	No feature selection	Feature selection

					Physics based Features	All Features	Selected features (wavelet coefficients)
Noise level	120dB	100.00%	100.00%	100.00%	100.00%	100.00%	**100.00%**
	110dB	97.71%	100.00%	98.86%	98.86%	98.86%	**100.00%**
	100dB	81.14%	86.29%	84.00%	92.00%	84.00%	**95.43%**
	90dB	44.00%	64.00%	72.00%	80.00%	72.00%	**86.29%**
	80dB	19.43%	34.86%	39.43%	53.71%	39.43%	**56.00%**

Further study also confirmed that full automation using deep learning could dramatically improve the accuracy of classification as compared to semi-automated methods.

Sub-Task 6.2. Decision-tree based human-machine decision making

Subtask 6.1 provided a comparative study of human machine interaction, ranging from physics-based methods to full automated methods. Clearly, the full automated methods are excellent tools for sensory data, in order to avoid any subjective interferences from human factor (such as variances due to different experience of inspectors). However, end-to-end full automation is a black box that could lead to false alarm in some case, where there are errors generated from data acquisition. Instead, expert experience and traceable and explainable data fusion could be a great complementary for informed decision making. In this case, ensemble of Resnet with decision tree could combine the coupled effects of human factors and machine-based automaton.

Decision tree is a highly interpretable machine learning method that splits the data according to the learned features and cutoff values. The relationship between the prediction y and the features x is:

$$y = \sum_{m=1}^{M} c_m I\{x \in R_m\} \quad (12)$$

where M is the number of subset, R_m is subset, I is the identical function that if $x \in R_m$ returns 1. c_m is the average of all training instances in subset R_m.

Take mechanical damage description for an example as plotted in **Fig. 10**, where the six-group data were virtually collected from the laboratory under scale factor from 1-5. Clearly, the data were associated with human factor, including personal training and field experience with respect to the topics. Note that due to different experience, such data may not reflect an identical trend as experienced pipe inspectors in pipe community who have worked in this field for years.

Fig. 55 showed the predicted results in terms of mechanical damage description from different weights. Data confirmed that the highest risk associated with human inputs, which could track back what are the major contributions from majority.

Real-world data will be collected in the coming periods and it will provide better representative for variances from different perspectives.

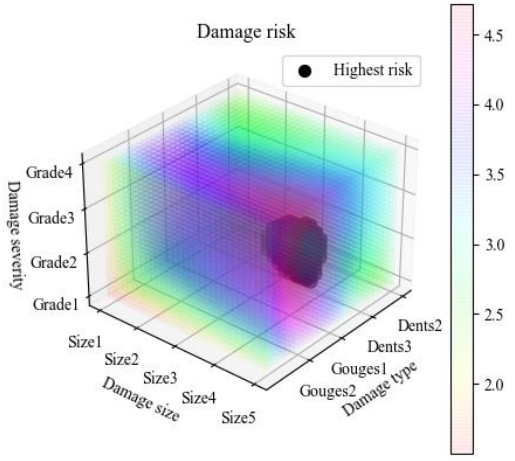


Fig. 10 Decision-tree based decision making for correlation among mechanical damage and risk

Sub-Task 6.3. Bayesian belief network-based human-machine decision making

6.3.1 Deep Bayesian Belief Network (DBBN)

Learning architecture of stochastic binary variables between layers $\boldsymbol{v}$ and $\boldsymbol{h}$ through the energy-based method using Boltzmann machine (Hinton et al. 2006; Hinton et al. 2012):

$$E(\boldsymbol{v},\boldsymbol{h}) = -\sum a_i v_i - \sum b_j h_j - \sum h'_j W v_i - \sum v'_i U v_i - \sum h'_j V h_j \quad (13)$$

where, a and b is the biases of the stochastic variables $\boldsymbol{v}$ and $\boldsymbol{h}$, respectively; W, U, V are the weights of each connection; and the joint states of v' and v *(or* h' and h) denote the adjacent connection of the variables within a layer. As illustrated in **Fig. 11**, the deep Bayesian belief network (Zhao et al., 2015; Chaturvedi et al., 2016) is the multiple layers neural networks. The model is effective to perform top-down and also the opposite bottom-up generative weights, which allows using back-propagation for fine-tuning for optimized discrimination/regression.

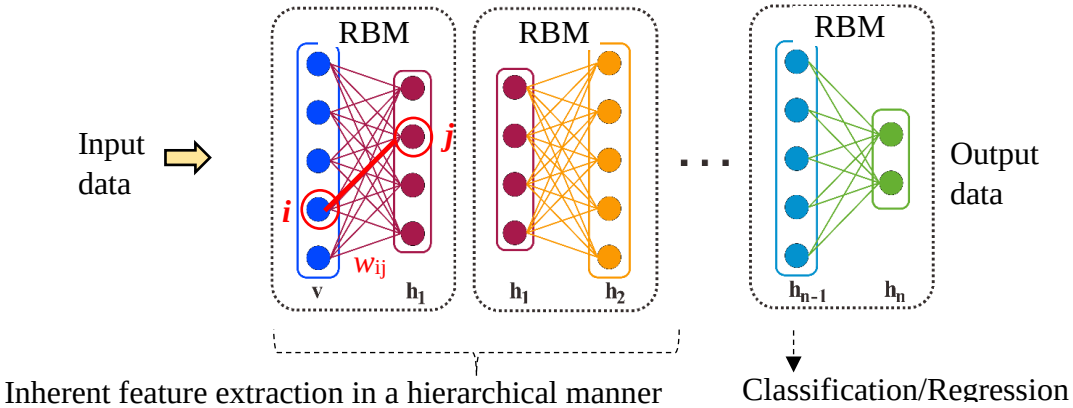


Fig. 11 Architecture of the DBBN

The DBBN are constructed using multiple restricted Boltzmann machines (RBMs) (Hinton et al. 2006; Hinton et al. 2012; Zhao et al., 2015). As schematically illustrated in **Fig. 8**, the architecture of the DBBN consists of undirected multiple levels of the RBMs, where the hierarchical architecture could allow automation of feature extraction from lower to upper layers, and the final layer could be constructed using different activation function for either classification or logistic regression of interest.

6.3.2 Concept of the RBM and its architectures

To avoid the complexity and difficulty in determining parameters in Boltzmann machine, as shown in Eqn. (5b), the RBM is developed by an undirected graphical machine without visible-visible or hidden-hidden connections (Mohamed et al. 2012). Take one unit of the RBM (e.g., RBM_1 in **Fig. 11**) as an example, the visible variable v_i, and the hidden variable, h_j, are connected and assigned by a weight, w_{ij}. The probability to the joint states of the visible and hidden vector is defined by the energy-based function (Hinton et al. 2006; Hinton et al. 2012)

$$p(\boldsymbol{v},\boldsymbol{h}) = \frac{e^{-E(\boldsymbol{v},\boldsymbol{h})}}{\sum_{i=1}^{V}\sum_{j=1}^{H} e^{-E(\boldsymbol{v},\boldsymbol{h})}} \quad (14)$$

where, $\sum$ is the summation over all visible and hidden variables and the $E()$ is the energy-based function. Consider structural data of interest could be binary data, such as black or white color in image recognition, or be more complex sensory information, three different data types (either in visible or hidden layer) are defined herein by (Hinton et al. 2010):

We will conduct a survey of pipeline owners, inspectors and other stakeholders in the coming periods to collect the critical information of human factors. These data will be used for calibration of the models and demonstration of the variances of human factors in decision making.

6.4. Summary of the research activities in the 2nd annual report

The major research activities are: (i) Task 4: Completed efforts on decoding uncertainty due to measurement noise, operational condition (temperature variances), and structural initial nonlinearity (integrated effects of weldments with allowable varying imperfection); aging effects toward data variances are in progress. (ii) Task 5: Completed efforts on decoding heterogeneous data sets using images and time-

series sensory data; (iii) Task 6: Development of learning for human-in-the-loop decision making, semi-automated human-machine decision making, and decision-tree based human-machine decision making. Survey and method calibration are in progress.

(e) Description of any Problems/Challenges

No problems are experienced during this report period

(f) Planned Activities for the Next Quarter

The planned activities for the next quarter are listed below:

- Variances due to aging effect of materials in Task 4 will be conducted in both numerical and experimental investigations.
- Variances due to pipe with and without coating in Task 4 will be conducted in both numerical and experimental investigations.
- A survey of pipeline owners, inspectors and other stakeholders will be conducted over the coming two quarter periods to collect the human factor associated critical information.
- Method calibration of decision-tree based human-machine decision making in Task 6 will be conducted.